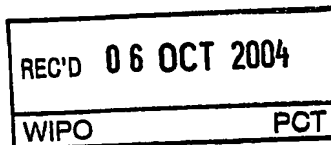




INVESTOR IN PEOPLE

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH
RULE 17.1(a) OR (b)



The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

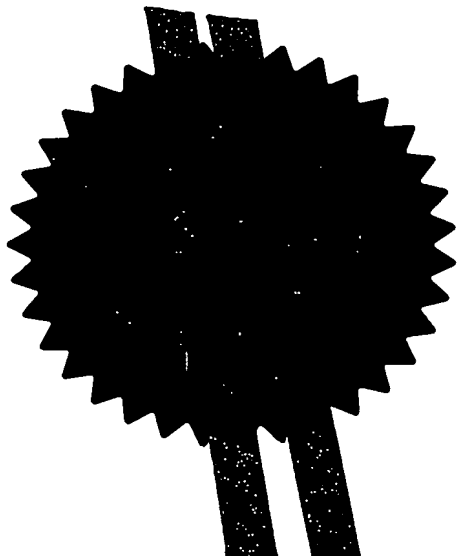
GB/04/3591

I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

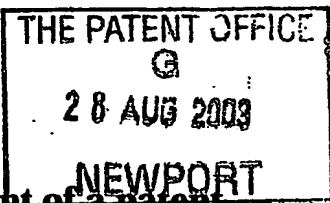
In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed

Dated 9 September 2004



1/77

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales

NP10 8QQ

1. Your reference

P342351/CPA/MCM

28AUG03 FB33248-6 802884
P01/7700 0.00-0320122.5

2. Patent application number

(The Patent Office will fill in this part)

28 AUG 2003

0320122.5

3. Full name, address and postcode of the or of each applicant (*underline all surnames*)

Albachem Limited
Elvingston Science Centre
by Gladsmuir
East Lothian
EH53 1EH

16 Charlotte Square
Edinburgh
EH2 4DF

086417970012

Patents ADP number (*if you know it*)

If the applicant is a corporate body, give the country/state of its incorporation

United Kingdom

4. Title of the invention

"Ligation Method"

5. Name of your agent (*if you have one*)

Murgitroyd & Company

"Address for service" in the United Kingdom to which all correspondence should be sent (*including the postcode*)

Scotland House
165-169 Scotland Street
Glasgow
G5 8PL

Patents ADP number (*if you know it*)

1198013

00001198015

6. If you are declaring priority from one or more earlier patent applications, give the country and the date of filing of the or of each of these earlier applications and (*if you know it*) the or each application number

Country

Priority application number
(*if you know it*)

Date of filing
(day / month / year)

7. If this application is divided or otherwise derived from an earlier UK application, give the number and the filing date of the earlier application

Number of earlier application

Date of filing
(day / month / year)

8. Is a statement of inventorship and of right to grant of a patent required in support of this request? (*Answer 'Yes' if:*

Yes

- a) any applicant named in part 3 is not an inventor, or
 - b) there is an inventor who is not named as an applicant, or
 - c) any named applicant is a corporate body.
- See note (d))

Patents Form 1/77

9. Enter the number of sheets for any of the following items you are filing with this form. Do not count copies of the same document

Continuation sheets of this form

Description

37

Claim(s)

7

Abstract

-

Drawing(s)

10 from 16

10. If you are also filing any of the following, state how many against each item.

Priority documents

Translations of priority documents

Statement of inventorship and right to grant of a patent (*Patents Form 7/77*)

Request for preliminary examination and search (*Patents Form 9/77*)

Request for substantive examination (*Patents Form 10/77*)

Any other documents
(please specify)

11.

I/We request the grant of a patent on the basis of this application.

Signature

Murgitroyd & Company

Date

27 August 2003.

12. Name and daytime telephone number of person to contact in the United Kingdom

Malcolm C Main

0141 307 8400

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

- If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.*
- Write your answers in capital letters using black ink or you may type them.*
- If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.*
- If you have answered 'Yes' Patents Form 7/77 will need to be filed.*
- Once you have filled in the form you must remember to sign and date it.*
- For details of the fee and ways to pay please contact the Patent Office.*

Ligation Method

Field of the Invention

This application relates to a method of ligating two or more molecules, for example, small organic molecules, labels, peptides etc. In particular it relates to a method of ligating a peptide, such as ligation of a synthetic peptide to a recombinant peptide.

Background to the Invention

Protein engineering methodologies have proven to be invaluable for generating protein based tools for application in basic research, diagnostics, drug discovery and as protein therapeutics. The ability to manipulate the primary structure of a protein in a controlled manner opens up many new possibilities in the biological and medical sciences. As a consequence, there is a concerted effort on developing methodologies for the site-specific modification of proteins and their subsequent application.

1
2 The two main approaches to generating proteins are
3 through recombinant methods or chemical synthesis.
4 To date, the two methods have proved to be
5 complementary; recombinant methodologies enable
6 proteins of any size to be generated but in general
7 they are restricted to the assembly of the
8 proteinogenic amino acids. Thus, in general, the
9 introduction of labels and probes into recombinant
10 proteins has to be implemented post-translationally
11 and does not allow modifications to the protein
12 backbone.

13
14 The most common methods for labelling a recombinant
15 protein use an amino or a thiol reactive version of
16 the label that will covalently react with a lysine
17 side chain / N^{α} amino group or a cysteine side chain
18 within the protein respectively. For such labelling
19 methods to be site-specific, an appropriate
20 derivative of the protein must be engineered to
21 contain a unique reactive functionality at the
22 position to be modified. This requires all the other
23 naturally occurring reactive functionalities within
24 the primary sequence to be removed through amino
25 acid mutagenesis. In the case of protein amino
26 functionalities, this is essentially impossible due
27 to the abundance of lysine residues and the presence
28 of the amino functionality at the N-terminus.
29 Likewise, for cysteine this process is laborious and
30 is often detrimental to the function of the protein.
31

1 The production of proteins having site-specific
2 modifications and/or labels is more readily
3 achievable using chemical synthesis methods. The
4 chemical synthesis of proteins, however, enables
5 multiple modifications to be incorporated into both
6 side-chain and backbone moieties of the protein in a
7 site-specific manner, but, in general, the maximum
8 size of sequence that can be synthesised and
9 isolated is circa 50 - 100 amino acids.

10

11 Protein Ligation

12 A further approach to the generation of proteins is
13 protein / peptide ligation. In this approach
14 mutually reactive chemical functionalities
15 (orthogonal to the chemistry of the naturally
16 occurring amino acids i.e. which react by mutually
17 exclusive chemistries compared to the reactions of
18 the reactive moieties of the naturally occurring
19 amino acids) are incorporated at the N- and C-
20 termini of unprotected polypeptide fragments such
21 that when they are mixed, they react in a
22 chemoselective manner to join the two sequences
23 together (Cotton GJ and Muir TW. Chem.Biol., 1999,
24 6, R247-R254). The principle of chemical ligation is
25 shown schematically in Figure 1.

26

27 A number of chemistries have been utilised for the
28 ligation of two synthetic peptides where a diverse
29 range of different chemical functionalities can be
30 incorporated into the termini of polypeptides using
31 solid phase peptide synthesis. These include the
32 reaction between a thioacid and bromo- alkyl to

1 form a thioester (Schnolzer M and Kent SBH, *Science*,
2 1992, 256, 221-225), reaction of an aldehyde with an
3 N-terminal cysteine or threonine to form
4 thiazolidine or oxazolidine respectively (Liu C-F
5 and Tam J P. *Proc. Natl. Acad. Sci. USA*, 1994, 91,
6 6584 - 6588), reaction between a hydrazide and an
7 aldehyde to form a hydrazone (Gaertner HF et al, et
8 al *Bioconj. Chem.*, 1992, 3, 262 - 268) reaction of
9 an aminoxy group and an aldehyde to form an oxime
10 (Rose K. *J. Am. Chem. Soc.*, 1994, 116, 30-33),
11 reaction of azides and aryl phosphines to form an
12 amide bond (Staudinger ligation) (Nilsson BL,
13 Kiessling LL, and Raines RT. *Org. Lett.*, 2001, 3, 9-
14 12, Kiick et al *Proc. Natl. Acad. Sci. USA*, 2002,
15 99, 19-24) , and the reaction of a peptide C-
16 terminal thioester and an N-terminal cysteine
17 peptide to form a native amide bond (Dawson et al.
18 *Science*, 1994, 266, 776) (Native chemical ligation
19 US6184344, EP 0832 096 B1). This method is an
20 extension of studies by Wieland and coworkers who
21 showed that the reaction of ValSPh and CysOH in
22 aqueous buffer yielded the dipeptide ValCysOH
23 (Wieland T et al, *Liebigs Ann. Chem.*, 1953, 583,
24 129-149).

25

26 Although the native chemical ligation method has
27 proved popular, it requires an N-terminal cysteine
28 and thus, if a cysteine is not present at the
29 appropriate position in the protein, a cysteine
30 needs to be introduced at the ligation site.
31 However, the introduction of extra thiol groups into
32 a protein sequence maybe detrimental to its

1 structure / function, especially since cysteine has
2 a propensity to form disulfide bonds which may
3 disrupt the folding pathway or compromise the
4 function of the folded protein.

5
6 As a consequence of the difficulties and problems
7 associated with known ligation techniques, the
8 ligation of two synthetic fragments generally only
9 enables proteins of circa 100 - 150 amino acids to
10 be chemically synthesised. Although larger proteins
11 have been synthesised by ligating together more than
12 two fragments, this has proved to be technically
13 difficult (Camarero et al. *J. Pept. Res.*, 1998, 54,
14 303-316, Canne LE et al, *J. Am. Chem. Soc.*, 1999,
15 121, 8720-8727).

16

17 Protein semi-synthesis

18

19 protein ligation technologies that enable both
20 synthetic and recombinantly derived protein
21 fragments to be joined together have been
22 described. This enables large proteins to be
23 constructed from combinations of synthetic and
24 recombinant fragments allowing proteins to be site-
25 specifically modified with both natural and
26 unnatural entities. By utilising such so-called
27 protein semi-synthesis, many different synthetic
28 moieties can be site-specifically incorporated at
29 multiple different sites within a target protein.

30

31 In order to utilise recombinant proteins in ligation
32 strategies the recombinant fragments must contain

1 the appropriate reactive functionalities to
2 facilitate ligation. One approach to introduce a
3 unique reactive functionality into a recombinant
4 protein has been through the periodate oxidation of
5 N-terminal serine containing sequences. Such
6 treatment converts the N-terminal serine into a
7 glyoxyl moiety, which contains an N-terminal
8 aldehyde. Synthetic hydrazide containing peptides
9 have then been ligated to the N-terminus of these
10 protein in a chemoselective manner through hydrazone
11 bond formation with the protein N-terminal aldehyde
12 group (Gaertner HF et al, et al Bioconj. Chem.,
13 1992, 3, 262 - 268, Gaertner HF, et al. *J. Biol.*
14 *Chem.*, 1994, 269, 7224-7230). Another approach has
15 been to generate recombinant proteins with N-
16 terminal cysteine residues. Synthetic peptides
17 containing C-terminal thioesters have then been
18 site-specifically attached to the N-terminus of
19 these proteins via amide bond formation in a manner
20 analogous to 'native chemical ligation' (Cotton GJ
21 and Muir TW. *Chem. Biol.*, 2000, 7, 253-261). However
22 as with the ligation of synthetic peptides using
23 native chemical ligation techniques, the technology
24 requires a cysteine to be introduced at the ligation
25 site if the primary sequence does not contain one at
26 the appropriate position.

27

28 Protein Splicing Techniques

29

30 Recently technologies have been developed which
31 enable recombinant proteins containing C-terminal
32 thioester groups to be generated. The C-terminal

1 thioester functionality provides a unique reactive
2 chemical group within the protein that can be
3 utilised for protein ligation. Recombinant C-
4 terminal thioester proteins are produced by
5 manipulating a naturally occurring biological
6 phenomenon known as protein splicing (Paulus H. Annu
7 Rev Biochem 2000, 69, 447-496). Protein splicing is
8 a post-translational process in which a precursor
9 protein undergoes a series of intramolecular
10 rearrangements which result in precise removal of an
11 internal region, referred to as an intein, and
12 ligation of the two flanking sequences, termed
13 exteins (Figure 2). While there are generally no
14 sequence requirements in either of the exteins,
15 inteins are characterised by several conserved
16 sequence motifs and well over a hundred members of
17 this protein domain family have now been identified.
18

19 The first step in protein splicing involves an N→S
20 (or N→O) acyl shift in which the N-extein unit is
21 transferred to the sidechain SH or OH group of a
22 conserved Cys/Ser/Thr residue, always located at the
23 immediate N-terminus of the intein. Insights into
24 this mechanism have led to the design of a number of
25 mutant inteins which can only promote the first step
26 of protein splicing (Chong et al Gene. 1997, 192,
27 271-281, (Noren et al., Angew. Chem. Int. Ed. Engl.,
28 2000, 39, 450-466). Proteins expressed as in frame
29 N-terminal fusions to one of these engineered
30 inteins can be cleaved by thiols via an
31 intermolecular transthioesterification reaction, to
32 generate the recombinant protein C-terminal

1 thioester derivative (Figure 3) (Chong et al *Gene*.
2 1997, 192, 271-281, (Noren et al., *Angew. Chem. Int.*
3 *Ed. Engl.*, 2000, 39, 450-466) (New England Biolabs
4 Impact System WO 00/18881, WO 0047751). Peptide
5 sequences containing an N-terminal cysteine residue
6 can then be specifically ligated to the C-termini of
7 such recombinant C-terminal thioester proteins (Muir
8 et al *Proc. Natl. Acad. Sci. USA.*, 1998, 95, 6705-
9 6710, Evans Jr et al. *Prot. Sci.*, 1998, 7, 2256-
10 2264) , in a procedure termed expressed protein
11 ligation (EPL) or intein-mediated protein ligation
12 (IPL). As with the previously described ligation
13 techniques, such an approach requires a cysteine to
14 be introduced at the ligation site if one is not
15 suitably positioned with the primary protein
16 sequence and thus is subject to the limitations and
17 associated with the problems of these approaches,
18 such as the potential problems associated with the
19 introduction of an extra thiol group into the
20 primary sequence.

21
22 The chemoselective ligation of N-terminal cysteine
23 containing peptides to C-terminal thioester
24 containing peptides, be they synthetic or
25 recombinant, is performed typically at slightly
26 basic pH and in the presence of a thiol cofactor.
27 The strategy also requires a cysteine to be
28 introduced at the ligation site, if one is not
29 suitably positioned within the primary sequence.
30 These requirements of this ligation approach have
31 the potential to alter the structure or function of

1 both the protein ligation product and the initial
2 reactants.

3

4 Protein labelling

5

6 Historically protein ligation means the joining
7 together of two peptide / protein fragments but this
8 is synonymous with protein labelling whereby the
9 label is a peptide or derivatised peptide. Equally
10 if a small non-peptidic synthetic molecule contains
11 the necessary reactive chemical functionality for
12 protein ligation, then ligation of the synthetic
13 molecule directly to either the N- or C- termini of
14 the protein affords site-specific labelling of the
15 protein. Thus technologies developed for the
16 ligation of protein fragments can also be used for
17 the direct labelling of either the N- or C- termini
18 of peptides or proteins in a site - specific manner
19 irrespective of their sequence.

20

21 Recombinant proteins containing N-terminal glyoxyl
22 functions (generated through periodate oxidation of
23 the corresponding N-terminal serine protein) have
24 been site-specific N-terminally labelled through
25 reaction with hydrazide or aminoxy derivatives of
26 the label (Geoghegan KF and Stroh JG. *Bioconj Chem.*,
27 1992, 3, 138-146, Alouni S et al. *Eur. J. Biochem.*,
28 1995, 227, 328 - 334). Also recombinant proteins
29 containing N-terminal cysteine residues have been N-
30 terminally labelled through reaction with thioester
31 containing labels, the label being the acyl
32 substituent of the thioester (Schuler B and Pannell

1 LK. *Bioconjug. Chem.*, 2002, 13, 1039-43) and
2 aldehyde (Zhao et al. *Bioconj. Chem.*, 1999, 10,
3 424-430) functionalities to form amides and
4 thiazolidines respectively.

5
6 Though a number of methods for ligation of proteins
7 exist each one has its potential drawbacks. There
8 is thus a need for novel ligation methodologies,
9 especially those that are compatible with both
10 synthetic and recombinant fragments, which will
11 complement the existing technologies and add another
12 string to the protein engineers' bow.

13 14 **Summary of the Invention**

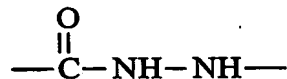
15
16 The present inventors have overcome a number of
17 problems associated with the prior art and have
18 developed a new method for ligating peptide
19 molecules which overcomes a number of the problems
20 of the prior art.

21
22 Accordingly, in a first aspect of the present
23 invention, there is provided a method of producing
24 an oligopeptide product, the method comprising the
25 steps:

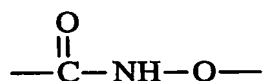
- 26 a) providing a first oligopeptide, the first
- 27 oligopeptide having a reactive moiety,
- 28 b) providing a second oligopeptide, the second
- 29 oligopeptide having a activated ester moiety
- 30 c) allowing the reactive moiety of the first
- 31 oligopeptide to react with the activated ester
- 32 moiety of the second oligopeptide to form an

oligopeptide product, in which the first and second oligopeptides are linked via a linking moiety having Formula I, Formula II or Formula III.

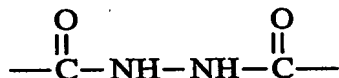
Formula I



Formula II



Formula III



In preferred embodiments, in step (c), where said oligopeptides are linked via a linking moiety having Formula II and where said activated ester moiety of step (b) is not a thioester, said activated ester is a terminal activated ester moiety.

In further preferred embodiments of the invention, said linking moieties are linked via a linking moiety having Formula I or Formula III.

Unless the context demands otherwise, the terms peptide, oligopeptide, polypeptide and protein are used interchangeably.

1 The activated ester moiety of the first oligopeptide
2 may be any suitable activated ester moiety, such as
3 a thioester moiety a phenolic ester moiety, an
4 hydroxysuccinimide moiety, or an O-acylisourea
5 moiety.

6
7 In preferred embodiments of the invention, the
8 activated ester moiety is a thioester moiety. Any
9 suitable thioester peptides may be used in the
10 present invention. In preferred embodiments, the
11 thioester is a thioester wherein the peptide is the
12 acyl substituent of the thioester.

13
14 Such thioester peptides may be synthetically or
15 recombinantly produced. The skilled person is well
16 aware of methods known in the art for generating
17 synthetic peptide thioesters. For example, synthetic
18 peptide thioesters may be produced via synthesis on
19 a resin that generates a C-terminal thioester upon
20 HF cleavage (Hojo et al, Bull. Chem. Soc. Jpn.,
21 1993, 66, 2700-2706). Further, the use of 'safety
22 catch' linkers has proved to be popular for
23 generating C-terminal thioesters through thiol
24 induced resin cleavage of the assembled peptide
25 (Shin Y et al, J. Am. Chem. Soc., 1999, 121, 11684-
26 11689).

27
28 Moreover, recently technologies have been developed
29 which enable recombinant C-terminal thioester
30 proteins to be generated. Recombinant C-terminal
31 thioester proteins may be produced by manipulating a
32 naturally occurring biological phenomenon known as

1 protein splicing. As described above, protein
2 splicing is a post-translational process in which a
3 precursor protein undergoes a series of
4 intramolecular rearrangements which result in
5 precise removal of an internal region, referred to
6 as an intein, and ligation of the two flanking
7 sequences, termed exteins.

8
9 As described above, a number of mutant inteins which
10 can only promote the first step of protein splicing
11 have been designed (Chong et al *Gene*. 1997, 192,
12 271-281, Noren et al., *Angew. Chem. Int. Ed. Engl.*,
13 2000, 39, 450-466). Proteins expressed as in frame
14 N-terminal fusions to one of these engineered
15 inteins can be cleaved by thiols via an
16 intermolecular transthioesterification reaction, to
17 generate the recombinant protein C-terminal
18 thioester derivative (Chong et al *Gene*. 1997, 192,
19 271-281, Noren et al., *Angew. Chem. Int. Ed. Engl.*,
20 2000, 39, 450-466) (New England Biolabs Impact
21 System WO 00/18881, WO 0047751). Such protein
22 thioesters may be used in the methods of the
23 invention (See Figure 3).

24
25 Accordingly, in a preferred aspect of the present
26 invention, in step (b), the second oligopeptide is
27 generated by thiol reagent induced cleavage of an
28 intein.

29
30 Accordingly, in a second aspect of the present
31 invention, there is provided a method of producing
32 an oligopeptide product, the method comprising the

1 steps:

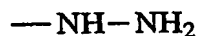
- 2 a) providing a first oligopeptide, the first
3 oligopeptide having a reactive moiety,
4 b) (i) providing a precursor oligopeptide
5 molecule, the precursor oligopeptide molecule
6 comprising a second oligopeptide fused N-terminally
7 to an intein domain
8 (ii) allowing thiol reagent dependent cleavage of
9 the precursor molecule to generate a second
10 oligopeptide molecule, said second oligopeptide
11 molecule having a thioester moiety at its C-terminus
12 c) allowing the reactive moiety of the first
13 oligopeptide to react with the second oligopeptide
14 molecule to form an oligopeptide product, in which
15 the first and second oligopeptides are linked via a
16 linking moiety having Formula I, II or III.

17

18 The reactive moiety of the first oligopeptide may be
19 any suitable reactive moiety. In preferred
20 embodiments of the invention, the reactive moiety is
21 a hydrazine moiety, an amino-oxy moiety or a
22 hydrazide moiety having general formula IV, V or VI
23 respectively.

24

25 Formula IV



26

27

28

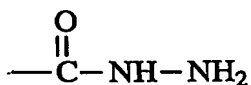
29 Formula V



30

31

1
2 Formula VI
3



4
5 For example, in a particular preferred embodiment,
6 the reactive moiety has Formula IV and, in the
7 oligopeptide product produced by the method of the
8 invention, the first and second oligopeptides are
9 linked via a linking moiety having Formula I.

10
11 In a further preferred embodiment, the reactive
12 moiety has Formula V and, in the oligopeptide
13 product produced by the method of the invention, the
14 first and second oligopeptides are linked via a
15 linking moiety having Formula II.

16
17 In another preferred embodiment, the reactive moiety
18 has Formula VI and, in the oligopeptide product
19 produced by the method of the invention, the first
20 and second oligopeptides are linked via a linking
21 moiety having Formula III.

22
23 As described above, the first oligopeptide comprises
24 a reactive moiety, which, in preferred embodiments,
25 may be a hydrazine moiety (e.g. Formula IV), an
26 amino-oxy moiety (e.g. Formula V) or an hydrazide
27 moiety (e.g. Formula VI).

28
29 Hydrazine, hydrazide or aminooxy containing
30 derivatives of synthetic oligopeptides may be
31 readily produced using known methods, for example,
32 solid phase synthesis techniques.

1

2 Further, the present inventors have also found that
3 proteins fused N-terminal to an intein domain can be
4 cleaved from the intein by hydrazine treatment in a
5 selective manner to liberate the desired protein as
6 its corresponding hydrazide derivative (for example,
7 see Figure 5).

8

9 Accordingly, in further preferred embodiments of the
10 invention, the first oligopeptide is generated by
11 reaction of hydrazine with an oligopeptide molecule
12 comprising the first oligopeptide fused N-terminal
13 to an intein domain.

14

15 Indeed the discovery that such protein hydrazides
16 may be produced by means of such a reaction forms an
17 independent aspect of the present invention.

18

19 Accordingly, a third aspect of the invention
20 provides a method of generating a protein hydrazide,
21 said method comprising the steps:

22 (a) providing an protein molecule comprising an
23 oligopeptide fused N-terminal to an intein domain,
24 (b) reacting said protein molecule with hydrazine,
25 such that the intein domain is cleaved from the
26 oligopeptide to generate a protein hydrazide.

27

28 Moreover, as well as using such a reaction to
29 generate a first oligopeptide having a hydrazide
30 moiety at its C-terminal, the first oligopeptide
31 thus being available for reaction with the second
32 oligopeptide having the activated ester moiety, the

1 present invention further extends to a "one-step"
2 process for ligating two peptides to generate an
3 oligopeptide product.

4

5 This may be achieved by reacting a suitable protein
6 linked N-terminal to an intein directly with a
7 polypeptide having a hydrazine, hydrazide or amino-
8 oxy moiety.

9

10 Accordingly, in a fourth aspect, the invention
11 provides a method of producing an oligopeptide
12 product, the method comprising the steps:

- 13 a) providing a first oligopeptide, the first
14 oligopeptide having a reactive moiety, wherein the
15 reactive moiety is a hydrazine moiety, a hydrazide
16 moiety or an amino-oxy moiety;
17 (i) providing a precursor oligopeptide molecule, the
18 precursor oligopeptide molecule comprising a second
19 oligopeptide fused N-terminally to an intein domain;
20 (c) allowing the reactive moiety of the first
21 oligopeptide to react with the precursor
22 oligopeptide molecule to form an oligopeptide
23 product, in which the first and second oligopeptides
24 are linked via a linking moiety having Formula I,
25 Formula II or Formula III.

26

27 The ligation technology of the present invention can
28 thus utilise both synthetic and recombinant proteins
29 and peptides. It thus enables the ligation of two or
30 more synthetic, two or more recombinant or a mixture
31 of one or more synthetic with one or more
32 recombinant peptides.

1
2 Moreover, as well as providing a novel method of
3 ligating peptides, the present invention may be used
4 for the labelling of synthetic or recombinant
5 peptides.

6
7 Accordingly, in a fifth aspect of the present
8 invention, there is provided a method of labelling
9 an oligopeptide, the method comprising the steps:
10 a) providing a label molecule, the label molecule
11 having a reactive moiety,
12 b) providing the oligopeptide, the oligopeptide
13 having an activated ester moiety
14 c) allowing the reactive moiety of the label
15 molecule to react with the activated ester moiety of
16 the oligopeptide to form the labelled oligopeptide,
17 in which the label molecule and the oligopeptide are
18 linked via a linking moiety having Formula I,
19 Formula II or Formula III as defined above,

20
21 In preferred embodiments, in step (c), where said
22 label molecule and the oligopeptide are linked via a
23 linking moiety having Formula II and where said
24 activated ester moiety of step (b) is not a
25 thioester, said activated ester is a terminal
26 activated ester moiety.

27
28 Alternatively, a label molecule having a terminal
29 activated ester moiety may be used to label an
30 oligopeptide having a reactive moiety. Thus, in a
31 sixth aspect of the invention, there is provided a
32 method of labelling an oligopeptide, the method

1 comprising the steps:

- 2 a) providing a label molecule, the label molecule
3 having an activated ester moiety of which the label
4 is the acyl substituent,
5 b) providing the oligopeptide, the oligopeptide
6 having a reactive moiety
7 c) allowing the activated ester moiety of the label
8 molecule to react with the reactive moiety of the
9 oligopeptide to form the labelled oligopeptide, in
10 which the label molecule and the oligopeptide are
11 linked via a linking moiety having Formula I,
12 Formula II or Formula III

13 wherein, in step (c), where said label molecule
14 and the oligopeptide are linked via a linking moiety
15 having Formula II and where said activated ester
16 moiety of step (b) is not a thioester, said
17 activated ester is a terminal activated ester
18 moiety.

19

20 As with the ligation technology, an oligopeptide
21 present as a precursor molecule linked to an intein
22 molecule may be labelled directly. Thus, a seventh
23 aspect of the present invention provides a method of
24 labelling an oligopeptide, the method comprising the
25 steps:

- 26 a) providing a label molecule, the label molecule
27 having a reactive moiety,
28 b) providing a precursor oligopeptide molecule,
29 the precursor oligopeptide molecule comprising an
30 oligopeptide fused N-terminally to an intein domain,
31 c) allowing the reactive moiety of the label
32 molecule to react with the precursor oligopeptide

1 molecule to form a labelled oligopeptide product, in
2 which the label molecule and the oligopeptide are
3 linked via a linking moiety having Formula I,
4 Formula II or Formula III as defined above.

5
6 The methods of the invention are particularly useful
7 in the ligation of peptides, in particular the
8 ligation of peptides, which, using conventional
9 ligation techniques, would require various
10 protecting groups. The inventors have shown that
11 the methods of the invention may be performed under
12 pH conditions in which only the reactive moieties
13 will react.

14
15 In preferred embodiments of the first to seventh
16 aspects of the invention, the method is performed at
17 a pH in the range pH 4.0 to pH 8.5, preferably pH
18 4.0 to 7.5, more preferably in the range pH 4.5 to
19 pH 7.0, most preferably in the range pH 5.5 to pH
20 6.5.

21
22 For example, the inventors have demonstrated that
23 synthetic peptide C-terminal thioesters specifically
24 react with hydrazine under aqueous conditions at pH
25 6.0 to form the corresponding peptide hydrazide.
26 This allows ligation methods as described herein to
27 be performed at pH 6.0, without the need for a
28 potentially harmful thiol cofactor (useful if either
29 fragment or final construct is thiol sensitive) and
30 does not lead to the introduction of potentially
31 reactive side-chain groups (such as a thiol) into
32 the protein. Similarly, the inventors have

1 demonstrated that synthetic peptide C-terminal
2 thioesters specifically react with hydroxylamine
3 under aqueous conditions at pH 6.0 and pH 6.8 to
4 form the corresponding peptide hydroxamic acid.
5

6 In an analogous fashion, peptides that contain
7 hydrazine, hydrazide or aminooxy groups can be
8 reacted with thioester derivatives of a label or a
9 peptide to afford site-specific labelling and
10 chemoselective ligation respectively (see, for
11 example, figures 4 and 5).
12

13 Furthermore, having demonstrated that recombinant
14 protein hydrazides can be generated by cleavage of
15 protein-intein fusions with hydrazine, the inventors
16 have shown that such protein hydrazides may be
17 ligated by reaction of the hydrazide moiety with
18 reactive groups other than activated ester moieties,
19 for example an aldehyde functionality, a ketone
20 functionality or an isocyanate functionality. This
21 aspect of the invention provides a further novel
22 method of ligating a recombinant peptide to a second
23 peptide or indeed a label.
24

25 Thus, an eighth aspect of the invention provides a
26 method of producing an oligopeptide product, the
27 method comprising the steps:

- 28 a) providing a first oligopeptide, the the first
- 29 oligopeptide having an aldehyde or ketone moiety,
- 30 b) providing a precursor oligopeptide molecule,
- 31 the precursor oligopeptide molecule comprising a
- 32 second oligopeptide fused N-terminally to an intein

1 domain,
2 c) reacting said precursor oligopeptide molecule
3 with hydrazine to generate an oligopeptide molecule
4 comprising an intermediate oligopeptide , said
5 intermediate oligopeptide having a C-terminal
6 hydrazide moiety,
7 d) allowing the aldehyde or ketone moiety of the
8 first oligopeptide to react with the hydrazide
9 moiety of the intermediate oligopeptide molecule to
10 form an oligopeptide product, in which first
11 oligopeptide and the second oligopeptide are linked
12 via a hydrazone linking moiety.
13
14 An example of this aspect is shown in Figure 6.
15
16 A ninth aspect of the invention provides a method of
17 labelling an oligopeptide, the method comprising the
18 steps:
19 a) providing a label molecule, the label molecule
20 having a aldehyde or ketone moiety,
21 b) providing a precursor oligopeptide molecule,
22 the precursor oligopeptide molecule comprising a
23 first oligopeptide fused N-terminally to an intein
24 domain,
25 c) reacting said precursor oligopeptide molecule
26 with hydrazine to generate an oligopeptide molecule
27 comprising an intermediate oligopeptide , said
28 intermediate oligopeptide having a terminal
29 hydrazide moiety,
30 d) allowing the aldehyde or ketone moiety of the
31 label molecule to react with the hydrazide moiety of
32 the intermediate oligopeptide molecule to form a

1 labelled oligopeptide product, in which the label
2 molecule and oligopeptide are linked via a hydrazone
3 linking moiety.

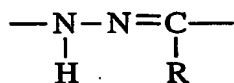
4

5 In preferred embodiments of the eighth and ninth
6 aspects of the invention, the hydrazone moiety has
7 Formula VII:

8

9

10



11

12 where R is H or any substituted or unsubstituted,
13 preferably unsubstituted, alkyl group.

14

15 In preferred aspects of the eighth and ninth aspects
16 of the invention, the method is performed at a pH in
17 the range pH 1.0 to pH 7.0, preferably pH 1.0 to pH
18 6.0, more preferably in the range pH 2.0 to pH 5.5,
19 most preferably in the range pH 2.0 to pH 4.5.

20

21 In a tenth aspect of the present invention, there is
22 provided an oligopeptide product produced using a
23 method of the invention.

24

25 In an eleventh aspect, there is provided a labelled
26 oligopeptide comprising an oligopeptide labelled
27 according to a method of the invention.

28

29 Preferred features of each aspect of the invention
30 are as for each of the other aspects mutatis
31 mutandis.

32

1 The invention will now be described further in the
2 following non-limiting examples with reference made
3 to the accompanying drawings in which:

4

5 Figure 1 illustrates schematically the general
6 principle of chemical ligation.

7

8 Figure 2 illustrates schematically the mechanism of
9 protein splicing.

10

11 Figure 3 illustrates the generation of recombinant
12 C-terminal thioester proteins.

13

14 Figure 4 illustrates ligation of protein and peptide
15 thioesters with hydrazine and aminooxy containing
16 entities, such as labels, peptides and proteins.

17

18 Figure 5 illustrates the generation of synthetic and
19 recombinant peptide hydrazides for ligation with
20 thioester containing molecules. Note the peptide or
21 label is the acyl substituent of the thioester.

22

23 Figure 6 illustrates the generation of recombinant
24 peptide hydrazides for ligation with aldehyde and
25 ketone containing molecules.

26

27 Figure 7 illustrates SDS-PAGE analysis of Grb2-SH2 -
28 GyrA - CBD (immobilised on chitin beads) treated
29 with DTT and MESNA. Molecular weight markers (lane
30 1); purified Grb2-SH2 - GyrA - CBD immobilised on
31 chitin beads (lane 4). Grb2-SH2 - GyrA - CBD treated
32 with 100 mM DTT (lanes 5 and 7) or 120 mM MESNA

1 (lanes 8 and 10). Both the whole reaction slurries
2 (lanes 5 and 8) and the reaction supernatants (lanes
3 7 and 10) were analysed.

4

5 Figure 8 illustrates SDS-PAGE analysis of Grb2-SH2 -
6 GyrA - CBD (immobilised on chitin beads) treated
7 with hydrazine. Molecular weight markers (lane 1);
8 Purified Grb2-SH2 - GyrA - CBD immobilised on chitin
9 beads after 20h treatment with phosphate buffer only
10 (lane 2). Grb2-SH2 - GyrA - CBD treated with 200 mM
11 hydrazine in phosphate buffer for 20 h. The whole
12 reaction slurries were analysed.

13

14 Figure 9 illustrates an ESMS spectrum of the C-
15 terminal hydrazide derivative of Grb2-SH2.

16

17 Figure 10 shows SDS-PAGE analysis of the reaction
18 between synthetic ketone containing peptide CH₃COCO-
19 myc with Grb2-SH2 - C-terminal hydrazide and
20 Cytochrome C. Molecular weight markers (lane 1);
21 Grb2-SH2 - C-terminal DTT thioester (lane 2).
22 Reaction between Grb2-SH2 - C-terminal hydrazide and
23 CH₃COCO-myc at time points t=0 h (lane 3), t=24 h
24 (lane 4), t= 48h (lane 5) and t= 72 h (lanes 6).
25 Reaction between Cytochrome C and CH₃COCO-myc at
26 time points t=0 h (lane 7), t=24 h (lane 8), t= 48h
27 (lane 9) and t= 72 h (lanes 10)

28

29

30 **Examples**

31

1 Example 1 -Protein ligation / site specific protein
2 labelling using the reaction of peptide / protein
3 thioesters with compounds containing hydrazine /
4 hydrazide or aminoxy functionalities.

5
6 Reaction of a peptide C-terminal thioester with
7 100mM hydrazine at pH 6.0
8 200 mM sodium phosphate buffer pH 6.0 containing
9 100mM hydrazine monohydrate (200 μ L) was added to a
10 model synthetic peptide α -thioester termed AS626p1A
11 (200 μ g) to yield a final peptide concentration of
12 317 μ M. AS626p1A has sequence ARTKQ TARK(Me)₃
13 STGGKAPRKQ LATKAARK-COS-(CH₂)₂-COOC₂H₅ (SEQ ID NO: 1)
14 wherein a single Alanine residue (which may be any
15 one of the Alanine residues of SEQ ID NO: 1) is
16 substituted by an Arginine residue. The reaction was
17 incubated at room temperature and monitored with
18 time by analytical reversed phase HPLC. Vydac C18
19 column (5 μ M, 0.46 x). Linear gradients of
20 acetonitrile water / 0.1% TFA were used to elute the
21 peptides at a flow rate of 1 mL min⁻¹. Individual
22 peptides eluting from the column were characterised
23 by electrospray mass spectrometry.

24

25 Reaction of a peptide C-terminal thioester with
26 100mM hydroxylamine at pH 6.0
27 200 mM sodium phosphate buffer pH 6.0 containing
28 100mM hydroxylamine hydrogen chloride (200 μ L) was
29 added to AS626p1A (200 μ g) to yield a final peptide
30 concentration of 317 μ M. The reaction was incubated
31 at room temperature and monitored with time by

1 analytical reversed phase HPLC. Vydac C18 column (5
2 μM , 0.46 x). Linear gradients of acetonitrile water
3 / 0.1% TFA were used to elute the peptides at a flow
4 rate of 1 mL min⁻¹. Individual peptides eluting from
5 the column were characterised by electrospray mass
6 spectrometry.

7

8 *Reaction of a peptide C-terminal thioester with 100*
9 *mM hydroxylamine at pH 6.8*

10 200 mM sodium phosphate buffer pH 6.8 containing
11 100mM hydroxylamine hydrogen chloride (200 μL) was
12 added to AS626p1A (200 μg) to yield a final peptide
13 concentration of 317 μM . The reaction was incubated
14 at room temperature and monitored with time by
15 analytical reversed phase HPLC. Vydac C18 column (5
16 μM , 0.46 x). Linear gradients of acetonitrile water
17 / 0.1% TFA were used to elute the peptides at a flow
18 rate of 1 mL min⁻¹. Individual peptides eluting from
19 the column were characterised by electrospray mass
20 spectrometry.

21

22 *Reaction of a peptide C-terminal thioester with 10mM*
23 *hydroxylamine at pH 6.8*

24 See above procedure.

25

26 *Reaction of a peptide C-terminal thioester with 10mM*
27 *hydroxylamine at pH 7.5*

28 See above procedure.

29

30 *Reaction of a peptide C-terminal thioester with 2mM*
31 *hydroxylamine at pH 7.5*

1 See above procedure.

2

3 Results

4 These examples demonstrate the novel strategy for
5 protein ligation / site specific protein labelling
6 of both synthetic and recombinant protein sequences
7 of the invention using the reaction of peptide /
8 protein C-terminal thioesters with compounds
9 containing hydrazine / hydrazide or aminoxy
10 functionalities.

11

12 As described above, a purified synthetic 27 amino
13 acid α -thioester peptide (the ethyl 3-
14 mercaptopropionate thioester derivative) was treated
15 with hydrazine and hydroxylamine under various
16 conditions (Table 1).

17

18 Treatment with 100 mM hydrazine at pH 6.0 formed a
19 peptide species that eluted earlier than the
20 starting thioester peptide as analysed by HPLC. This
21 material was identified as the expected peptide
22 hydrazide by ESMS: observed mass = 3054 Da, expected
23 (av. isotope comp) 3053 Da. The reaction of the
24 peptide C-terminal thioester with hydrazine to form
25 the peptide hydrazide was monitored with time by
26 reverse phase HPLC. Only the desired material was
27 formed with no side product formation even after 3
28 days. The stability of the peptide hydrazide, under
29 the reaction conditions, indicates that the reaction
30 occurs at the C-terminal thioester moiety and is
31 chemoselective in nature. It also highlights the
32 applicability of this reaction for protein ligation

1 and labelling. (2 h 70% conversion , 4h 95%
2 conversion)

3

4 To ascertain whether aminooxy containing compounds
5 chemoselectively react with peptide / protein C-
6 terminal thioesters, to afford protein ligation and
7 site-specific labelling, a synthetic C-terminal
8 thioester peptide was treated with hydroxylamine
9 under various conditions (Table 1).

10

11 A purified synthetic 27 amino acid C-terminal
12 thioester peptide (ethyl 3-mercaptopropionate
13 thioester, observed mass 3155 Da) was incubated at
14 room temperature with different hydroxylamine
15 concentrations in aqueous buffers of varying pH. In
16 all cases the peptide C-terminal thioester reacted
17 to form a single product that eluted earlier than
18 the starting thioester peptide as analysed by
19 reverse phase HPLC. This material corresponds to the
20 expected hydroxamic acid peptide as determined by
21 ESMS: observed mass = 3052 Da, expected (av. isotope
22 comp) 3054 Da. The kinetics of the reaction were
23 monitored using reverse phase HPLC. The peptide C-
24 terminal thioester is converted to the corresponding
25 peptide hydroxamic acid in a clean fashion with no
26 side-product formation. As expected increasing the
27 pH of the reaction buffer accelerates the rate of
28 reaction. With 100mM NH_2OH on moving from pH 6.0 to
29 pH 6.8 the percentage product formation after 1h
30 increases from 25% to 91%. The rate of reaction with
31 100 mM NH_2OH pH 6.0 is comparable with 10 mM NH_2OH at
32 pH 6.8.

1
2 The rate of reaction of the peptide C-terminal
3 thioester with hydroxylamine, to form the
4 corresponding hydroxamic acid, increases with
5 increasing pH and decreases with decreasing NH_2OH
6 concentrations. To identify conditions of pH and
7 reactant concentration suitable for peptide /
8 protein labelling and ligation, the labelling was
9 performed under increasing pH and decreasing NH_2OH
10 concentrations.

11
12 The reaction with 10 mM was 83% complete after 4h at
13 pH 6.8, while at pH 7.5 it was 83% complete after
14 2h. On further decreasing the NH_2OH concentration to
15 2 mM the reaction rate at pH 7.5 decreased markedly,
16 70% of the starting peptide α -thioester being
17 converted to the corresponding hydroxamic acid after
18 8hrs. It was noted that a small amount of a side-
19 product corresponding in mass to the peptide acid
20 was formed during the reaction. Presumably this is
21 formed by a competing hydrolysis side reaction at pH
22 7.5, which was not observed with 10 mM NH_2OH at pH
23 7.5 due to the faster reaction at this higher
24 reactant concentration.

Reactant	Concentration	pH	Percentage product formation with time				
			1hr	2hr	4hr	8hr	72hr
NH ₂ NH ₂	100 mM	6.0	-	70	100		
NH ₂ OH	100 mM	6.0	25	48.1	76.3	-	100
NH ₂ OH	100 mM	6.8	91	100			
NH ₂ OH	10 mM	6.8	26	-	83	100	
NH ₂ OH	10 mM	7.5	-	82.7	100	100	
NH ₂ OH	2 mM	7.5	11.2	17	38	70	80*

Table 1

*All starting material has reacted with 80% conversion to the desired product and ~20% to the hydrolysis side-product.

Example 2- Generation of recombinant C-terminal hydrazide proteins through the selective cleavage of protein - intein fusions with hydrazine, and their subsequent use in ligation / labelling reactions.

To investigate (i) the ability to generate recombinant C-terminal hydrazide proteins through the selective cleavage of protein - intein fusions with hydrazine, and (ii) their subsequent use in ligation / labelling reactions, the SH2 domain of the adapter protein Grb2 was chosen as a model system.

Sequence of human Grb2 SH2 domain

HPW FFGKIPRAKA EEMLSKQRHD GAFLIRESES APGDFSLSVK
FGNDVQHFKV LRDGAGKYFL WVKFNSLNE LVDYHRSTSV
SRNQQIFLRD IEQVPQQPT

1 *Expression of Grb2-SH2 domain - GyrA intein fusion.*

2 The DNA sequence encoding the SH2 domain of
3 human Grb2 appended at its C-terminus with an extra
4 glycine residue was cloned into the pTXB1 expression
5 plasmid (NEB). This vector pTXB1_{Grb2-SH2 (Gly)} encodes
6 for a fusion protein whereby the SH2 domain of Grb2
7 is linked via a glycine residue to the N-terminus of
8 the GyrA intein, which is in turn fused to the N-
9 terminus of a chitin binding domain region (CBD).
10 *E. coli* cells were transformed with this plasmid and
11 grown in LB medium to mid log phase and protein
12 expression induced for 4h at 37°C with 0.5 mM IPTG.
13 After centrifugation the cells were re-suspended in
14 lysis buffer (0.1 mM EDTA, 250 mM NaCl, 5% glycerol,
15 1 mM PMSF, 25 mM HEPES, pH 7.4) and lysed by
16 sonication. The soluble fraction was loaded onto a
17 chitin column pre- equilibrated in lysis buffer. The
18 column was then washed with wash buffer (1 mM EDTA,
19 250 mM NaCl, 0.1% Triton-X 100, , 25 mM HEPES, pH
20 7.0) to yield purified Grb2-SH2 - GyrA-CBD
21 immobilised on chitin beads (Figure 7).

22
23 *Generation of Grb2-SH2 C-terminal thioesters by*
24 *thiol induced cleavage of the Grb2-SH2 - GyrA intein*
25 *fusion.*

26 To ascertain that the intein domain within the
27 protein was functional the fusion protein was
28 exposed to thiols to assess the extent of cleavage
29 via transthioesterification. Chitin beads containing
30 immobilised Grb2-SH2 - GyrA-CBD were equilibrated
31 into 200 mM NaCl, 200 mM phosphate buffer pH 7.4.
32 Dithiothreitol (DTT) or 2-mercaptoethanesulfonic

1 acid (MESNA) were then added to the beads in 200 mM
2 NaCl, 200 mM phosphate buffer pH 7.4 to give a 50%
3 slurry with a final thiol concentration of 100 mM or
4 120 mM respectively. The mixtures were then rocked
5 at room temperature and aliquots analysed by SDS-
6 PAGE. After 48 hours the supernatants from the
7 reactions were isolated and subsequently analysed by
8 HPLC and ESMS.

9 Treatment of Grb2-SH2 - GyrA intein - CBD
10 fusion with both DTT and MESNA resulted in cleavage
11 of the fusion protein into two protein species
12 (Figure 7). The molecular size of the two fragments
13 corresponds to that of the Grb2 - SH2 and the GyrA -
14 intein fusion, indicative that cleavage has taken
15 place at the SH2 - intein junction. Cleavage of the
16 precursor fusion protein liberated the SH2 domain
17 into the supernatant while the GyrA intein-CBD
18 portion remained immobilized on the chitin beads.
19 After cleavage with both DTT or MESNA, ESMS analysis
20 of the supernatants confirmed that the Grb2-SH2 was
21 generated as either the expected DTT or MESNA C-
22 terminal thioester derivatives respectively.

23 Expected mass of Grb2-SH2 DTT - C-terminal
24 thioester = 12173.9 Da; observed mass 12173.5 Da.
25 Expected mass of Grb2-SH2 MESNA - C-terminal
26 thioester = 12162.0 Da; observed mass 12163.0 Da.

27

28 *Generation of Grb2-SH2 C-terminal hydrazide by*
29 *hydrazine induced cleavage of the Grb2-SH2 - GyrA*
30 *intein fusion.*

31

1 The thioester linkage between Grb2-SH2 and the
2 GyrA intein in the precursor fusion protein is
3 expected to be cleaved with hydrazine, the
4 chemoselective reaction of hydrazine, at the
5 thioester moiety, liberating Grb2-SH2 domain into
6 the supernatant as its corresponding C-terminal
7 hydrazide derivative. Chitin beads containing
8 immobilised Grb2-SH2 - GyrA-CBD were therefore
9 equilibrated into 200 mM NaCl, 200 mM phosphate
10 buffer pH 7.4 and hydrazine monohydrate added in the
11 same buffer to give a 50% slurry with a final
12 hydrazine concentration of 200 mM. The mixture was
13 then rocked at room temperature and analysed by SDS-
14 PAGE (Figure 8). After 20 hours the supernatant was
15 removed and analysed by HPLC and ESMS.

16 Treatment of Grb2-SH2 - GyrA intein - CBD
17 fusion with hydrazine resulted in cleavage of the
18 fusion protein into two species. The molecular size
19 of the two fragments as analysed by SDS-PAGE
20 corresponded to Grb2 - SH2 and the GyrA - intein
21 fusion, indicative that cleavage has taken place at
22 the unique thioester linkage between the SH2 -
23 intein domains. Cleavage of the precursor fusion
24 protein liberates the SH2 domain into the
25 supernatant while the GyrA intein-CBD portion
26 remained immobilized on the chitin beads. HPLC and
27 ESMS analysis of the cleavage supernatant confirmed
28 that a single protein species was generated that
29 corresponds to the C-terminal hydrazide derivative
30 of Grb2-SH2. Expected mass of Grb2-SH2 C-terminal
31 hydrazide = 12051.7 Da; observed mass 12053.0 Da.
32 (Figure 9)

1
2 After 20 h of reaction Grb2-SH2 C-terminal hydrazide
3 was isolated from the supernatant using RPHPLC and
4 lyophilised.

5
6 *Ligation of aldehyde and ketone containing peptides*
7 *and labels to recombinant C-terminal hydrazide*
8 *containing proteins.*

9
10 It was anticipated that recombinant protein C-
11 terminal hydrazides, generated by hydrazine
12 treatment of the corresponding intein fusion
13 precursor, can be site-specifically modified by
14 chemoselective ligation with aldehyde and ketone
15 containing peptides and labels. To demonstrate such
16 an approach the ability of a synthetic ketone
17 containing peptide to ligate with the Grb2-SH2 C-
18 terminal hydrazide generated above was investigated.
19 A synthetic peptide corresponding to the c-myc
20 epitope sequence was synthesised GEQKLISEEDL-NH₂
21 whereby pyruvic acid was coupled to the amino
22 terminus of the peptide as the last step of the
23 assembly. This peptide (designated CH₃COCO-myc) was
24 purified to > 95% purity by RPHPLC and lyophilised
25 (ESMS expected monoisotopic mass 1328.6 Da; observed
26 mass 1328.6 Da).

27 A sample of CH₃COCO-myc peptide was dissolved
28 in 100 mM sodium acetate buffer pH 4.5 to give a 4
29 mM peptide concentration. This peptide solution (100
30 µL) was then added to an aliquot of lyophilised
31 Grb2-SH2 C-terminal hydrazide protein (~ 250 µg) and
32 the reaction monitored by SDS-PAGE (Figure 10) As a

1 control CH₃COCO-myc was also incubated with
2 Cytochrome C, a protein of similar same size to
3 Grb2-SH2 but absent of a hydrazide functionality.

4 SDS-PAGE analysis showed that CH₃COCO-myc
5 peptide has indeed ligated with Grb2-SH2 C-terminal
6 hydrazide as indicated by the conversion of Grb2-
7 SH2 C-terminal hydrazide into a protein species of
8 a higher molecular weight (approximately 1000-2000
9 Da higher). The reaction was virtually complete
10 after 24 h and the reaction product appeared to be
11 stable. On the other hand there was no observable
12 change to Cytochrome C with time i.e no ligation,
13 establishing that the ligation reaction was
14 occurring at the C-terminal hydrazide functionality
15 of Grb2-SH2.

16 After 96 h of reaction the product from the
17 Grb2-SH2 ligation reaction was isolated by HPLC and
18 characterised by ESMS. Chemosselective ligation of
19 CH₃COCO-myc to Grb2-SH2 C-terminal hydrazide via
20 hydrazone bond formation would give a product of
21 expected mass 13363.7 Da. The observed product mass
22 was 13364.1 Da indicating that the desired ligation
23 product had been formed.

24

25 In summary, the present invention provides novel
26 methods of protein ligation that enable both
27 synthetic and recombinantly derived protein
28 fragments to be efficiently joined together in a
29 regioselective manner. This thus enables large
30 proteins to be constructed from combinations of
31 synthetic and recombinant fragments and allows
32 proteins of any size to be site-specifically

1 modified in an unprecedented manner. This is of
2 major importance for biological and biomedical
3 science and drug discovery when one considers that
4 the ~ 30,000 human genes yield hundreds of thousands
5 of different protein species through post-
6 translational modification. Such post-
7 translationally modified proteins cannot be accessed
8 through current recombinant technologies.

9
10 The application of such protein ligation techniques
11 may be used for protein based tools, protein
12 therapeutics and in *de novo* design and may open up
13 many new avenues in biological and biomedical
14 sciences that have hitherto not been possible.

15
16 All documents referred to in this specification are
17 herein incorporated by reference. Various
18 modifications and variations to the described
19 embodiments of the inventions will be apparent to
20 those skilled in the art without departing from the
21 scope and spirit of the invention. Although the
22 invention has been described in connection with
23 specific preferred embodiments, it should be
24 understood that the invention as claimed should not
25 be unduly limited to such specific embodiments.
26 Indeed, various modifications of the described modes
27 of carrying out the invention which are obvious to
28 those skilled in the art are intended to be covered
29 by the present invention.

30
31

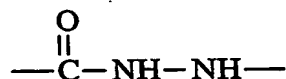
1 **Claims**

2

- 3 1. A method of producing an oligopeptide product,
 4 the method comprising the steps:
 5 a) providing a first oligopeptide, the first
 6 oligopeptide having a reactive moiety,
 7 b) providing a second oligopeptide, the second
 8 oligopeptide having a activated ester moiety
 9 c) allowing the reactive moiety of the first
 10 oligopeptide to react with the activated ester
 11 moiety of the second oligopeptide to form an
 12 oligopeptide product, in which the first and second
 13 oligopeptides are linked via a linking moiety having
 14 Formula I, Formula II or Formula III.

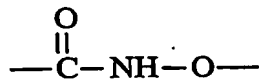
15

16 Formula I



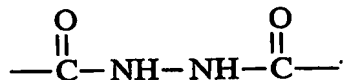
17

18 Formula II



19

20 Formula III



21

22

23

- 24 2. The method according to claim 1 wherein the
 25 terminal activated ester moiety is a thioester
 26 wherein the peptide is the acyl substituent of

1 the thioester.

2

3 3. The method according to claim 2, wherein said
4 second polypeptide is generated by thiol reagent
5 dependent cleavage of a precursor molecule, said
6 precursor molecule comprising a second oligopeptide
7 fused N-terminally to an intein domain.

8

9 4. A method of producing an oligopeptide product,
10 the method comprising the steps:

11 a) providing a first oligopeptide, the first
12 oligopeptide having a reactive moiety,
13 (i) providing a precursor oligopeptide molecule, the
14 precursor oligopeptide molecule comprising a second
15 oligopeptide fused N-terminally to an intein domain
16 (ii) allowing thiol reagent dependent cleavage of
17 the precursor molecule to generate a second
18 oligopeptide molecule, said second oligopeptide
19 molecule having a thioester moiety at its C-
20 terminus,
21 c) allowing the reactive moiety of the first
22 oligopeptide to react with the second oligopeptide
23 molecule to form an oligopeptide product, in which
24 the first and second oligopeptides are linked via a
25 linking moiety having Formula I, II or III.

26

27 5. The method according to any one of the preceding
28 claims wherein the reactive moiety is a hydrazine
29 moiety, a hydrazide moiety or an aminooxy moiety.

30

31 6. The method according to claim 5, wherein said
32 first oligopeptide is produced by reaction of

1 hydrazine with a precursor molecule, said
2 precursor molecule comprising a precursor
3 oligopeptide fused N-terminally to an intein
4 domain via a thioester moiety.
5

6 7. A method of producing an oligopeptide product,
7 said method comprising the steps:
8 a) providing a first oligopeptide, the first
9 oligopeptide having a reactive moiety, wherein
10 the reactive moiety is a hydrazine moiety, a
11 hydrazide moiety or an amino-oxy moiety;
12 (i) providing a precursor oligopeptide molecule,
13 the precursor oligopeptide molecule comprising a
14 second oligopeptide fused N-terminally to an
15 intein domain;
16 (c) allowing the reactive moiety of the first
17 oligopeptide to react with the precursor
18 oligopeptide molecule to form an oligopeptide
19 product, in which the first and second
20 oligopeptides are linked via a linking moiety
21 having Formula I, Formula II or Formula III.
22

23 8. The method according to any one of the preceding
24 claims, wherein the first oligopeptide or the
25 second oligopeptide is a recombinant oligopeptide
26 and the other of the the first oligopeptide and
27 the second oligopeptide is a synthetic
28 polypeptide.
29

30 9. The method according to any one of claims 1 to 7,
31 wherein the first oligopeptide and the second

1 oligopeptide are recombinant oligopeptides.

2

3 10. The method according to any one of claims 1 to
4 7, wherein the first oligopeptide and the second
5 oligopeptide are synthetic oligopeptides.

6

7 11. A method of generating a protein hydrazide,
8 said method comprising the steps:
9 (a) providing a protein molecule comprising an
10 oligopeptide fused N-terminal to an intein
11 domain,
12 (b) reacting said protein molecule with
13 hydrazine, such that the intein domain is cleaved
14 from the oligopeptide to generate a protein
15 hydrazide.

16

17 12. The method according to any one of the
18 preceding claims wherein the method is performed
19 at a pH in the range pH 5.5 to 7.5.

20

21 13. A method of producing an oligopeptide product,
22 the method comprising the steps:
23 a) providing a first oligopeptide, the the first
24 oligopeptide having an aldehyde or ketone moiety,
25 b) providing a precursor oligopeptide molecule,
26 the precursor oligopeptide molecule comprising a
27 second oligopeptide fused N-terminally to an
28 intein domain,
29 c) reacting said precursor oligopeptide molecule
30 with hydrazine to generate an oligopeptide
31 molecule comprising an intermediate oligopeptide
32 , said intermediate oligopeptide having a

- 1 terminal hydrazide moiety,
2 d) allowing the aldehyde or ketone moiety of the
3 first oligopeptide to react with the hydrazide
4 moiety of the intermediate oligopeptide molecule
5 to form an oligopeptide product, in which first
6 oligopeptide and the second oligopeptide are
7 linked via a hydrazone linking moiety.
8
- 9 14. An oligopeptide product produced by the method
10 of any one of the preceding claims.
11
- 12 15. A method of labelling an oligopeptide, the
13 method comprising the steps:
14 a) providing a label molecule, the label molecule
15 having a reactive moiety,
16 b) providing the oligopeptide, the oligopeptide
17 having a activated ester moiety
18 c) allowing the reactive moiety of the label
19 molecule to react with the activated ester moiety
20 of the oligopeptide to form the labelled
21 oligopeptide, in which the label molecule and the
22 oligopeptide are linked via a linking moiety
23 having Formula I, Formula II or Formula III.
24
- 25 16. The method according to claim 15, wherein in
26 step (c), where said label molecule and the
27 oligopeptide are linked via a linking moiety
28 having Formula II and where said activated ester
29 moiety of step (b) is not a thioester, said
30 activated ester is a terminal activated ester
31 moiety.
32

1 17. A method of labelling an oligopeptide, the
2 method comprising the steps:
3 a) providing a label molecule, the label molecule
4 having an activated ester moiety of which the
5 label is the acyl substituent,
6 b) providing the oligopeptide, the oligopeptide
7 having a reactive moiety
8 c) allowing the activated ester moiety of the
9 label molecule to react with the reactive moiety
10 of the oligopeptide to form the labelled
11 oligopeptide, in which the label molecule and the
12 oligopeptide are linked via a linking moiety
13 having Formula I, Formula II or Formula III,
14 wherein, in step (c), where said label molecule
15 and the oligopeptide are linked via a linking
16 moiety having Formula II and where said activated
17 ester moiety of step (b) is not a thioester, said
18 activated ester is a terminal activated ester
19 moiety.

20
21 18. A method of labelling an oligopeptide, the
22 method comprising the steps:
23 a) providing a label molecule, the label molecule
24 having a reactive moiety,
25 b) providing a precursor oligopeptide molecule,
26 the precursor oligopeptide molecule comprising an
27 oligopeptide fused N-terminally to an intein
28 domain,
29 c) allowing the reactive moiety of the label
30 molecule to react with the precursor oligopeptide
31 molecule to form a labelled oligopeptide product,
32 in which the label molecule and the oligopeptide

1 are linked via a linking moiety having Formula I,
2 Formula II or Formula III as defined above.

3

4 19. The method according to any one of claims 15 to
5 18 wherein the method is performed at a pH in the
6 range pH 5.5 to pH 7.5.

7

8 20. A method of labelling an oligopeptide, the
9 method comprising the steps:
10 a) providing a label molecule, the label molecule
11 having a aldehyde or ketone moiety,
12 b) providing a precursor oligopeptide molecule,
13 the precursor oligopeptide molecule comprising a
14 first oligopeptide fused N-terminally to an
15 intein domain,
16 c) reacting said precursor oligopeptide molecule
17 with hydrazine to generate an oligopeptide
18 molecule comprising an intermediate oligopeptide,
19 said intermediate oligopeptide having a terminal
20 hydrazide moiety,
21 d) allowing the aldehyde or ketone moiety of the
22 label molecule to react with the hydrazide moiety
23 of the intermediate oligopeptide molecule to form
24 a labelled oligopeptide product, in which the
25 label molecule and oligopeptide are linked via a
26 hydrazone linking moiety.

27

28 21. A labelled oligopeptide produced by the method
29 of any one of claims 15 to 20.

30

31

Mutually reactive groups

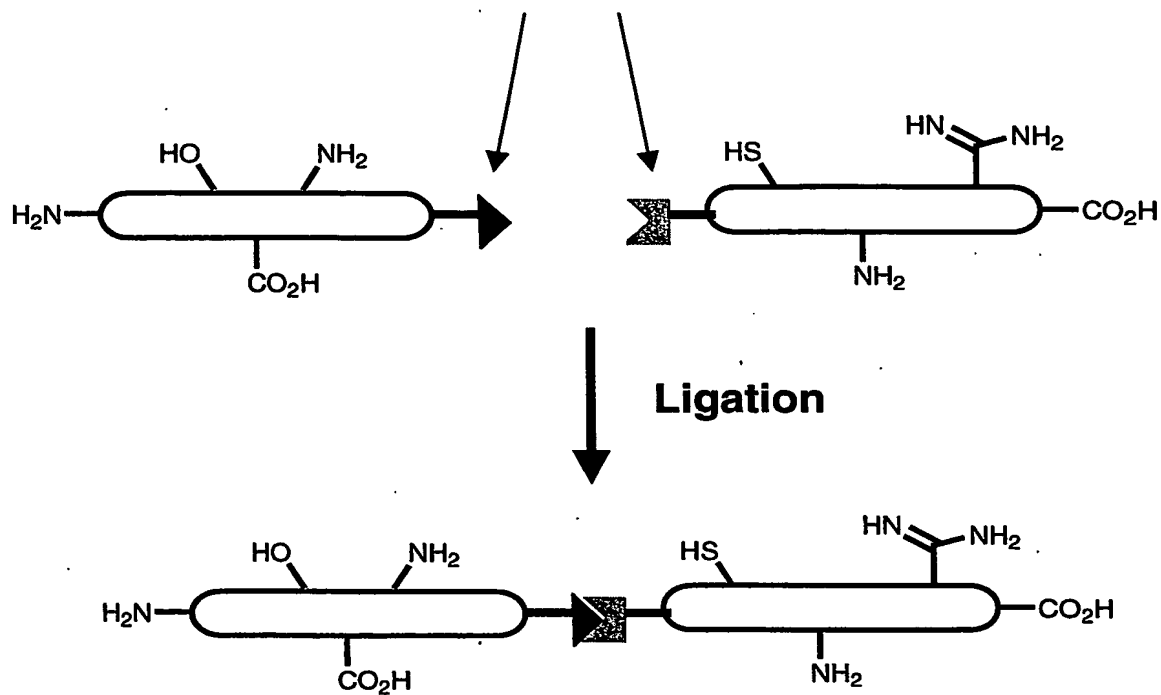


Figure 1 General principle of chemical ligation.

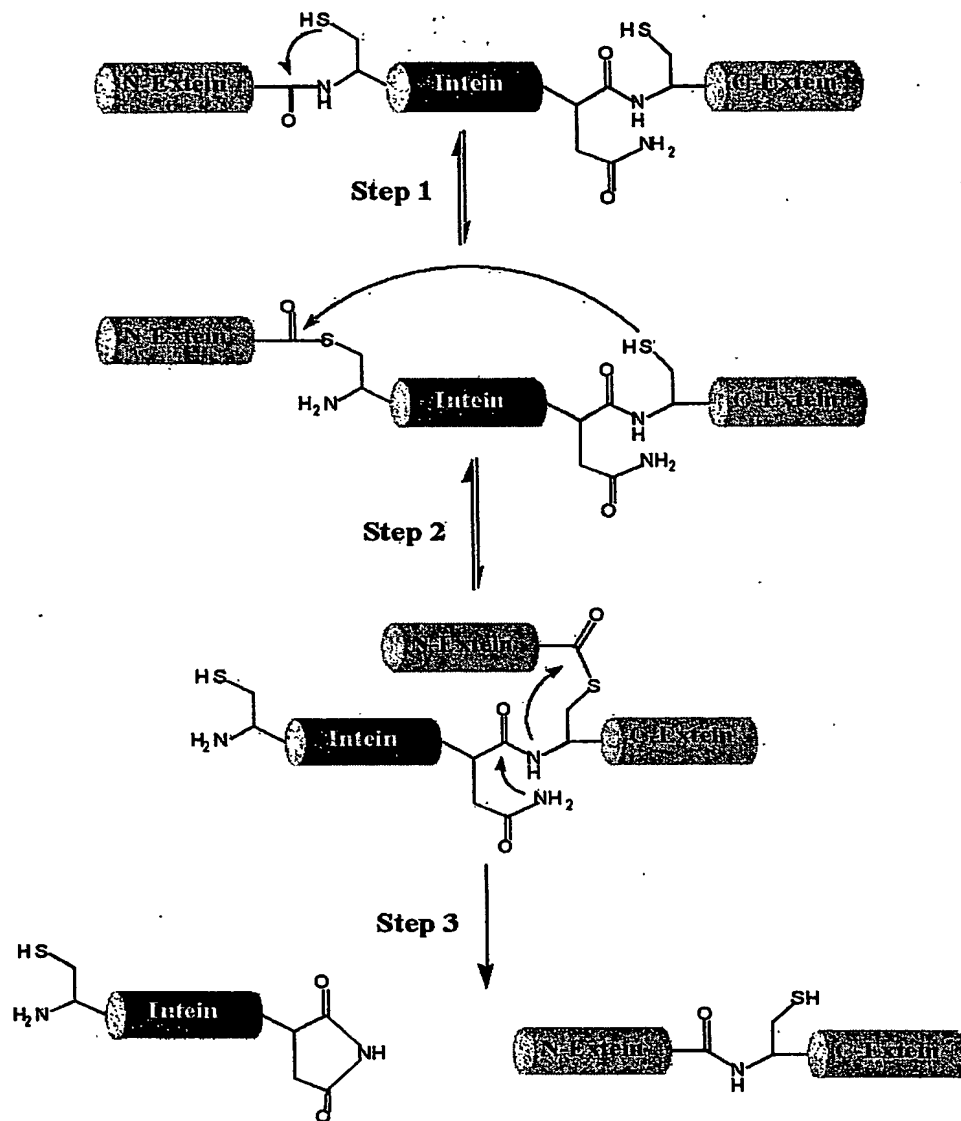


Figure 2 Mechanism of protein splicing

**Clone Gene into Engineered
Intein Expression Vector**

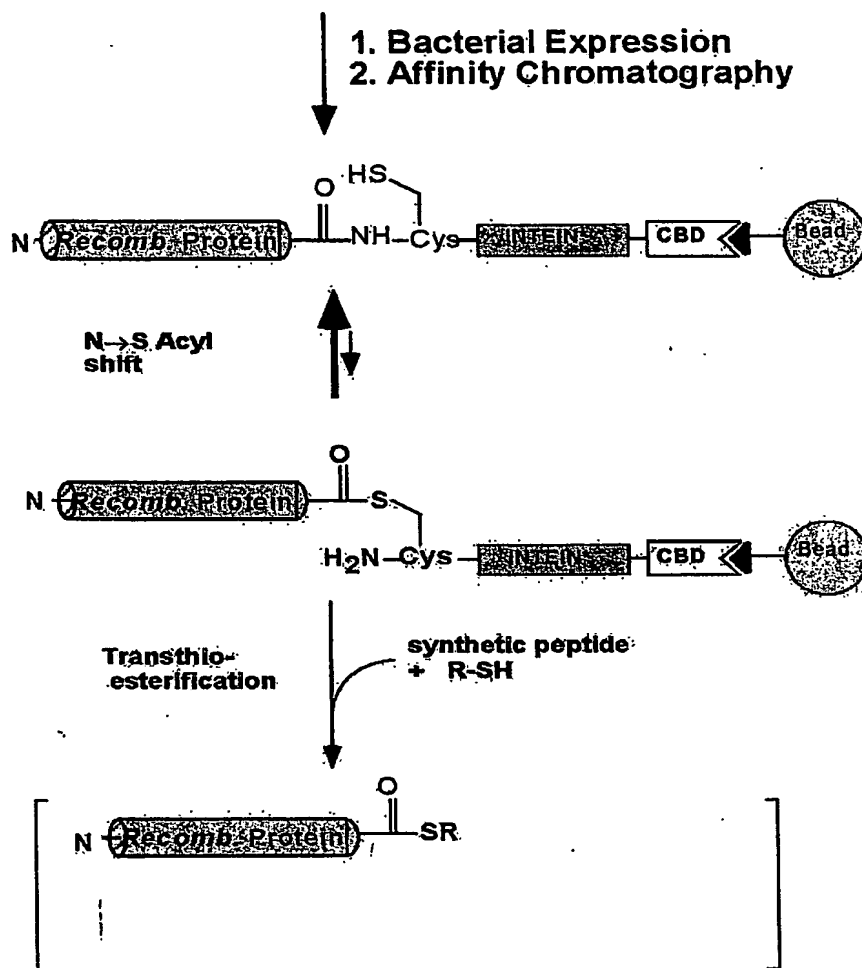


Figure 3 Generation of Recombinant C-terminal Thioester Proteins

Synthetic or recombinant peptide / protein -thioester

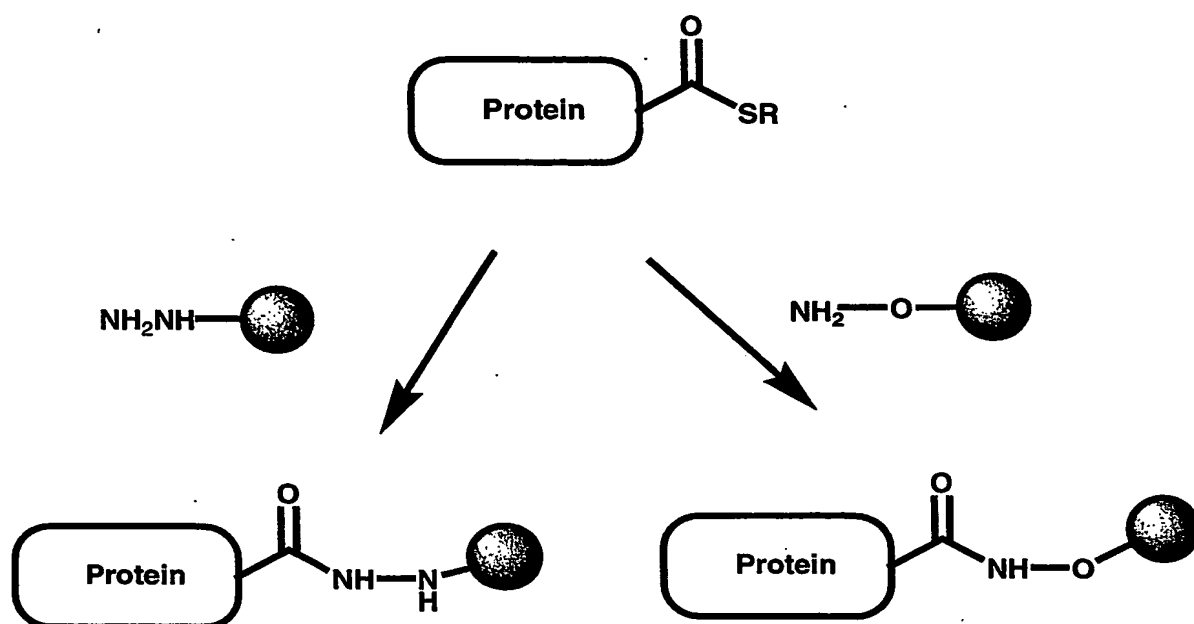


Figure 4 Ligation of protein and peptide thioesters with hydrazine and aminooxy containing entities such as labels, peptides and proteins.

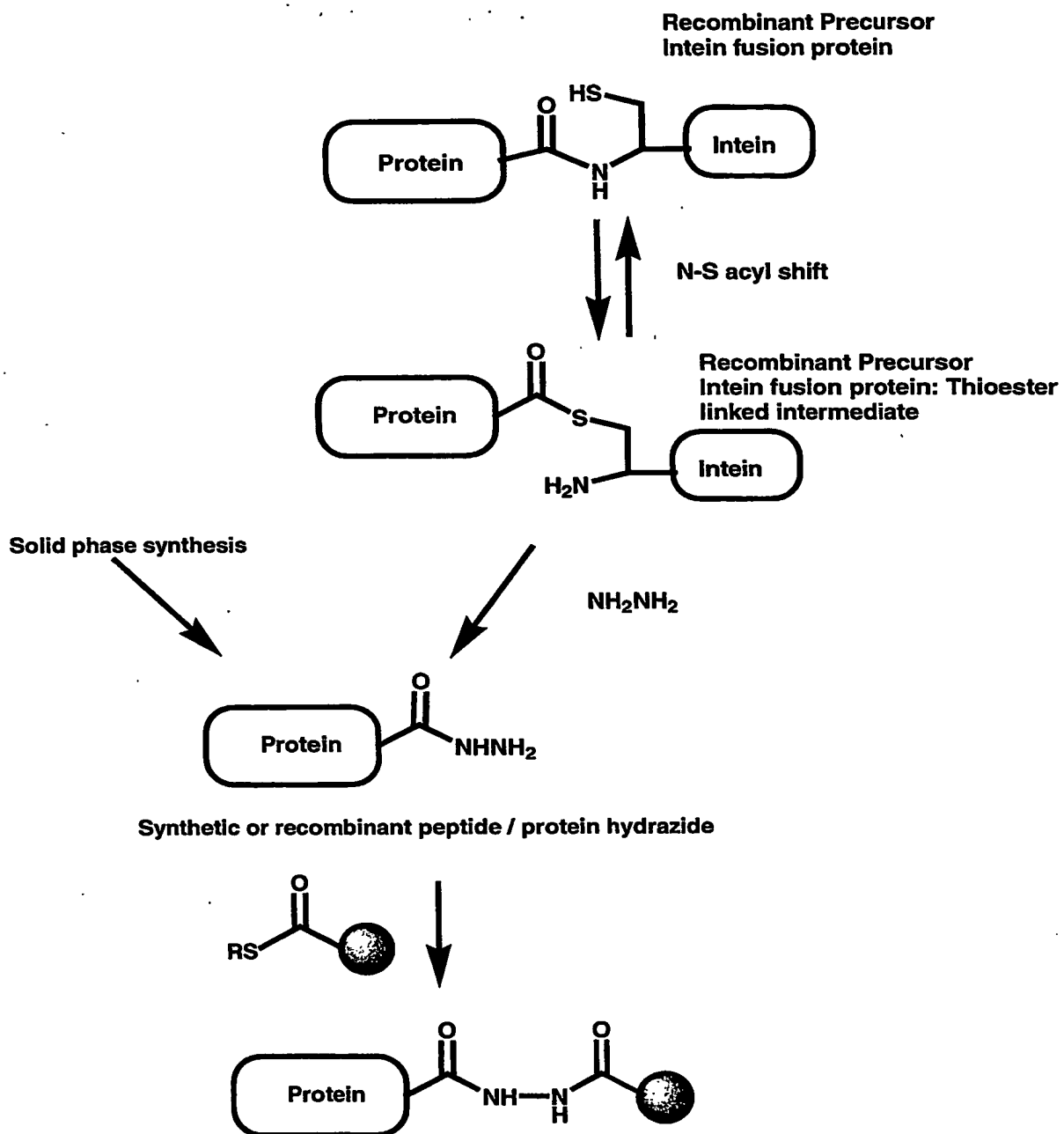


Figure 5 Generation of synthetic and recombinant peptide hydrazides for ligation with thioester containing molecules

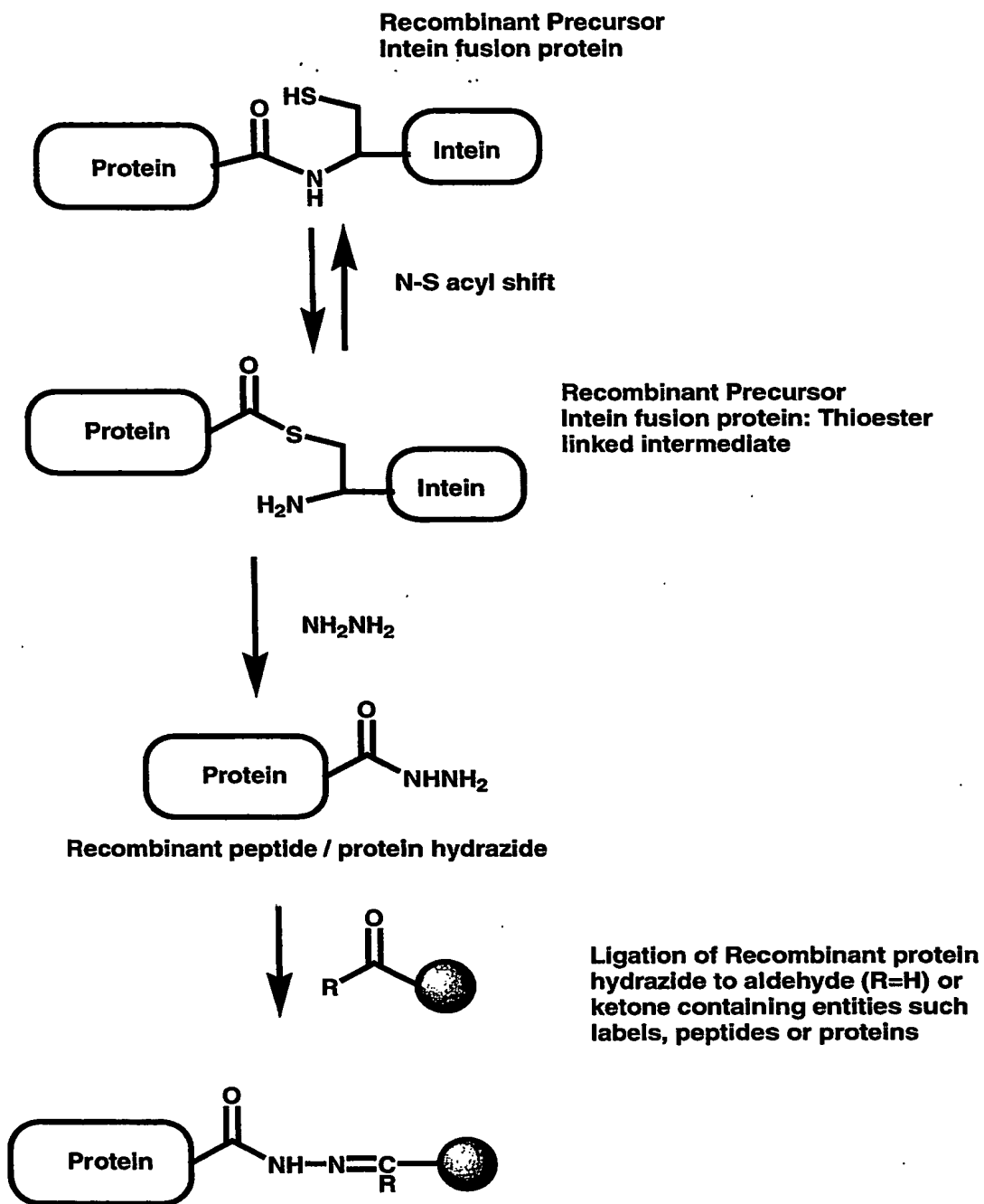


Figure 6 Generation of recombinant peptide hydrazides for ligation with aldehyde and ketone containing molecules

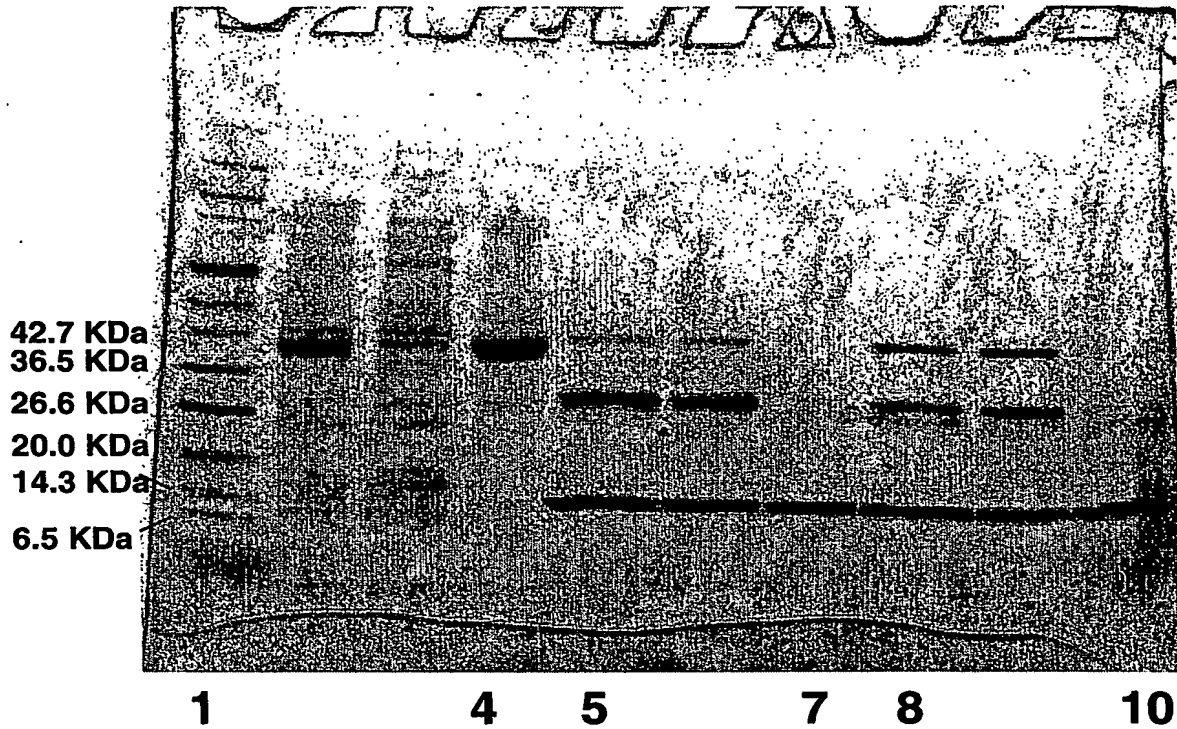


Figure 7. SDS-PAGE analysis of Grb2-SH2 – GyrA – CBD (immobilised on chitin beads) treated with DTT and MESNA. Molecular weight markers (lane 1); purified Grb2-SH2 – GyrA – CBD immobilised on chitin beads (lane 4). Grb2-SH2 – GyrA – CBD treated with 100 mM DTT (lanes 5 and 7) or 120 mM MESNA (lanes 8 and 10). Both the whole reaction slurries (lanes 5 and 8) and the reaction supernatants (lanes 7 and 10) were analysed.

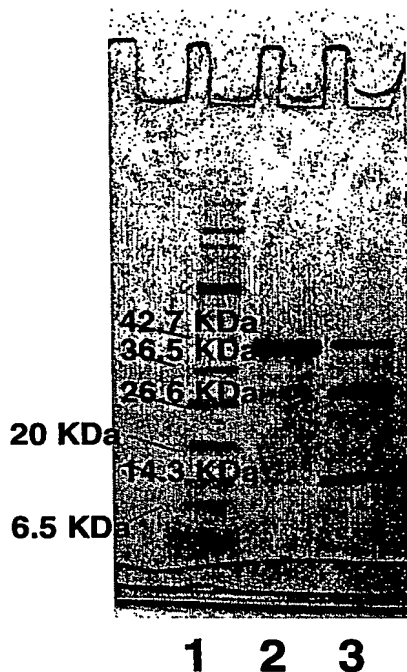


Figure 8. SDS-PAGE analysis of Grb2-SH2 – GyrA – CBD (immobilised on chitin beads) treated with hydrazine. Molecular weight markers (lane 1); Purified Grb2-SH2 – GyrA – CBD immobilised on chitin beads after 20h treatment with phosphate buffer only (lane 2). Grb2-SH2 – GyrA – CBD treated with 200 mM hydrazine in phosphate buffer for 20 h. The whole reaction slurries were analysed.

9/10

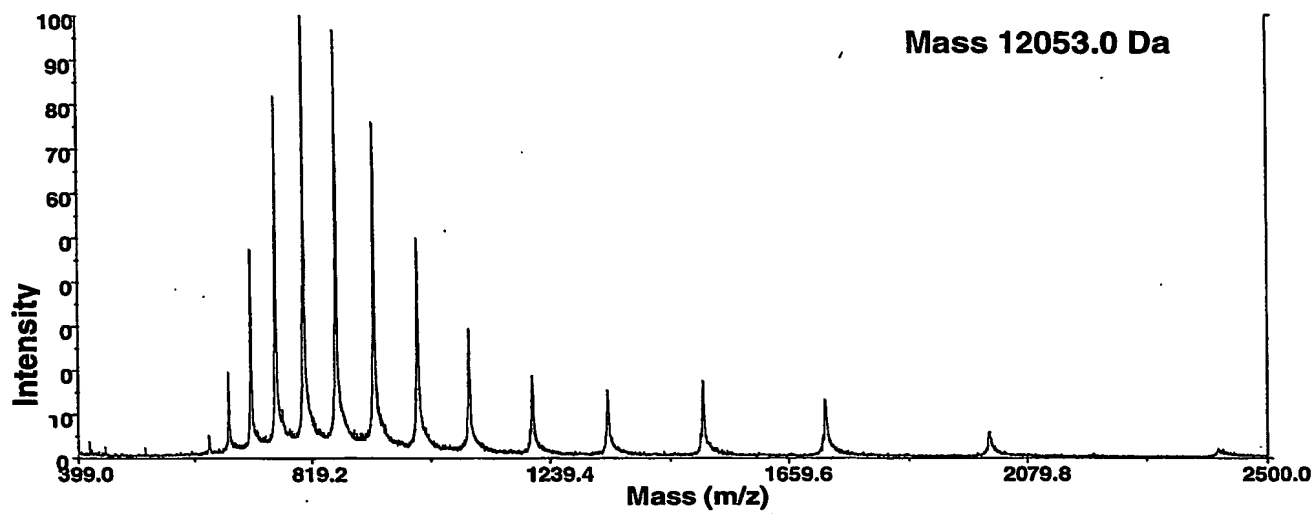


Figure 9. ESMS spectrum of the C-terminal hydrazide derivative of Grb2-SH2

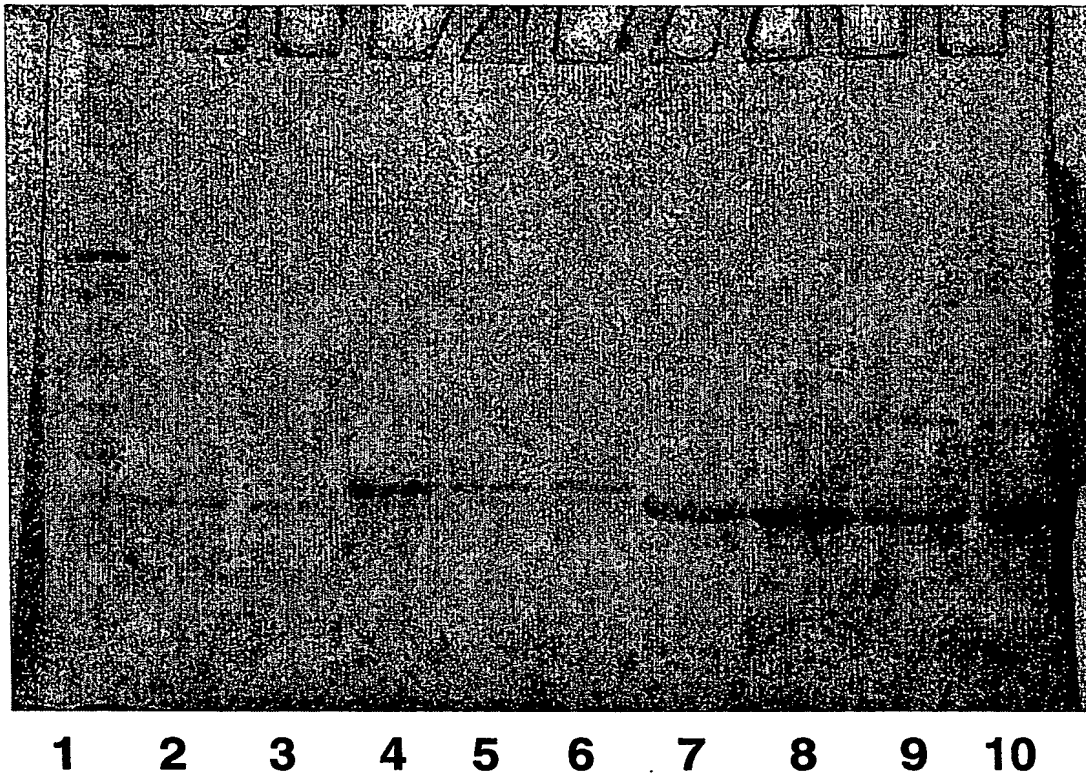


Figure 10. SDS-PAGE analysis of the reaction between synthetic ketone containing peptide CH₃COCO-myc with Grb2-SH2 – C-terminal hydrazide and Cytochrome C. Molecular weight markers (lane 1); Grb2-SH2 – C-terminal DTT thioester (lane 2). Reaction between Grb2-SH2 – C-terminal hydrazide and CH₃COCO-myc at time points t=0 h (lane 3), t=24 h (lane 4), t= 48h (lane 5) and t= 72 h (lanes 6). Reaction between Cytochrome C and CH₃COCO-myc at time points t=0 h (lane 7), t=24 h (lane 8), t= 48h (lane 9) and t= 72 h (lanes 10).